Pakistan J. Zool., vol. 45(5), pp. 1185-1190, 2013

Genetic Heterogeneity and Gene Diversity at *ABO* and *Rh* Loci in the Human Population of Southern Punjab, Pakistan

Sajid Malik[1]* and Muhammad Amin-ud-Din[2]
[1]*Human Genetics Program, Department of Animal Sciences, Quaid-i-Azam University, Islamabad 45320, Pakistan*
[2]*University of Education, DG Khan Campus, Kangan Road, DG Khan, Pakistan*

Abstract.- Southern Punjab in Pakistan is a heterogeneous assemblage of human populations which have diverse origins. Various waves of pre- and post-partition (1947) migrations, urbanization and high inbreeding coefficient have profoundly influenced the diversity and substructuring of this multiethnic and multilinguistic population. In order to get an insight into the genetic structure of Southern Punjab population, we have employed classical immuno-genetic markers of *ABO* and *Rh* loci. Five administrative districts, *i.e.*, Dera Ghazi Khan, Muzaffargarh, Multan, Bahawalpur, and Liaquatpur, were included in this study and the phenotypic and genotypic data of approximately 60,000 subjects was assembled. Various genetic parameters including maximum likelihood estimates of allelic frequencies, Hardy-Weinberg Equilibrium, locus specific and combined heterozygosities, and phenotypic and allelic variances, were investigated to understand population dynamics. Affinities between recruited populations were measured through tests of homogeneity, and gene diversity and coefficient of gene differentiation was computed in sub- and total-population. These analyses established that Southern Punjab population is markedly distinct from the upper Punjab/Punjab populations for the genetic systems studied. Additionally, population stratification and substructuring were significantly higher in Southern Punjab compared to upper Punjab/Punjab, depicting the likely role of micro-evolutionary forces shaping the dynamics of this interesting population. This pilot study would help understand the genetic structure of Southern Punjab population and to find its affinities with other Pakistani populations.

Key words: *ABO* and *Rh* genes, gene diversity, genetic heterogeneity, Southern Punjab.

INTRODUCTION

Southern Punjab is the southernmost region of the Punjab, Pakistan, comprising Dera Ghazi Khan (DG Khan), Multan, Muzaffargarh, Rajanpur, Bahawalpur and Rahimyar Khan districts. It is a predominantly Saraiki speaking region, and Indus and Chenab rivers contribute to its agriculture based economy. In Southern Punjab, as in other parts of Punjab and Sindh, migration has long played a key role in shaping the size and distribution of the population. There have been several waves of pre- and post-partition (1947) migrations. According to the most accepted tradition, Aryans, Baloch and Pathans settled this area in different influxes. Then, there was Arabian and Central Asian component of immigrants which also had permanent settlements. Furthermore, the settlers coming from different parts of the Indian subcontinent during and after 1947 largely mixed up with one of the allied ethnic group of this region (Abdullah, 1973; Mian and Farooq, 1999). Additionally, the local and micro-scale migrations also influenced the population structure. For instance, the migrants groups from Indian subcontinent, together with the growing number of rural people displaced by agriculture modernization and mechanization, have contributed to the substantial increase in the levels of urbanization (Hasan and Raza, 2009). Finally, the unique socio-cultural norms, extended family sizes, and endogamous and consanguineous marriage pattern resulted in further isolation and fragmentation within this multiethnic population. All these factors have acted as strong demographic forces shaping the Southern Punjab population. This situation makes Southern Punjab a very interesting region to explore the dynamics of its population.

In order to get an insight into the genetic structure of Southern Punjab population we have employed *ABO* and *Rh* blood loci. These immune-genetic markers are known to be excellent genetic systems with stable transmission, consistent

* Corresponding author: malik@qau.edu.pk
0030-9923/2013/0005-1185 $ 8.00/0

expression patterns and full penetrances. Additionally, their distributions have been worked out in extensively in human populations all over the world. It has become apparent that there is heterogeneity, not only between the continents and regions within them, but also on a microscale among neighboring counties and cities (Shami and Rasmuson, 1994). For the present study, five representative Southern Punjab populations were recruited which include DG Khan, Muzaffargarh, Multan, Bahawalpur and Liaquatpur (Rahim Yar Khan). The genetic heterogeneity and gene diversity analyses of *ABO* and *Rh* genes for Southern Punjab populations are being first time reported in this communication.

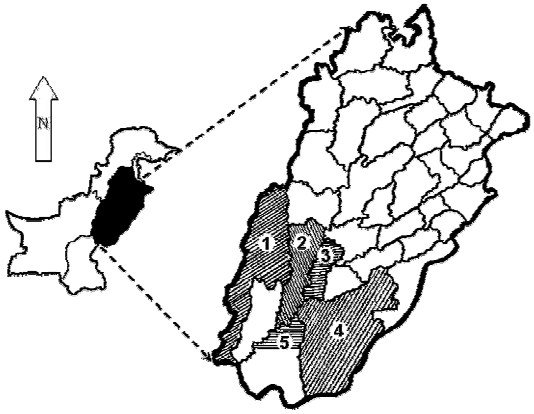

Fig. 1 Map of Pakistan with set-in Punjab province highlighting five recruited Southern Punjab populations. 1, DG Khan; 2, Muzaffargarh; 3, Multan; 4, Bhawalpur; 5, Liaquatpur. Districts Rajanpur and Rahim Yar Khan/Sadiqabad are left unfilled, as no substantial data on blood types is available for these populations.

SUBJECTS AND METHODS

Phenotypic data of 59,296 subjects from five Southern Punjab populations was assembled (Fig. 1, Table I). For DG Khan and Muzaffargarh Districts, data of 8,953 and 11,764 subjects, respectively, was collected from the visitors/volunteers of the respective District Headquarters Hospital. For Multan, Bahawalpur and Liaquatpur Districts, data was extracted from the previous reports (Mian and Bhutta, 1993; Khichi *et al*., 2000; Rehman *et al*., 2005).

From the phenotypic record of each population, maximum likelihood estimates of the allele frequencies were obtained and Hardy-Weinberg equilibrium (HWE) was tested (Mather, 1964; Silva, 2002). For each population, homogeneity was tested for the data acquired in consecutive years and with respect to gender (Neel and Schull, 1954). Z-test was employed to estimate the variances in phenotypes and to check the significance of heterogeneity of blood group proportions among the studied populations (Garstman, 2008). This gave an initial impression of similarities between populations. Subsequently, for a meaningful grouping of populations and to identify uniform clusters among them, homogeneity tests were conducted for *ABO* and *Rh* gene frequencies (Neel and Schull, 1954). Published data on the adjoining populations were utilized for the comparisons. To understand gene diversity, individual and combined heterozygosities at *ABO* and *Rh* loci were measured (Nei, 1975). The degree of differentiation among the five populations according to ABO/Rh polymorphic systems was calculated (Nei, 1975), and the results were displayed in unweighted pair group method with arithmetic mean (UPGMA) dendrogram (Sneath and Sokal, 1973). Low differentiation values between individual categories depicted more affinities and *vice versa.*

RESULTS

Basic demographic parameters, sample size and the phenotypic distributions of ABO and Rh blood groups of five Southern Punjab populations is provided in Table I. In the overall data, blood type O was predominant (35.68%) at the ABO system followed by B, A and AB blood types with the percentages of 35.58, 22.46, and 6.28, respectively. With respect to Rh system, Rh^- blood type was 5.22%

In the whole data, the maximum likelihood estimates of frequencies of alleles p[A], q[B], r[O], and Rh[d] were 0.1566, 0.2386, 0.6048, and 0.2285, respectively (Table II). Three populations (*i.e.*, Muzaffargarh, Bahawalpur, and Liaquatpur), were significantly deviating from the HWE test assumptions.

Table I.- Phenotypic distribution of ABO and Rh blood types in five populations of Southern Punjab.

Population / District	Area (Km^2)	Population 1998 *	n	Blood phenotype						Study year
				A	B	AB	O	Rh+	Rh-	
DG Khan	3,814	1,151,236	8,953	2,174	2,615	575	3,589	8,460	493	1996-97 (Present study)
Muzaffargarh	2,377	982,866	11,764	3,089	4,525	903	3,247	10,975	789	1997-98 (Present study)
Multan city	304	1,500,617	2,674	619	981	274	800	2,546	128	1988-91 (Mian and Bhutta, 1993)
Bahawalpur	2,372	806,580	34,516	7,147	12,525	1,942	12,902	32,963	1,553	2000 (Khichi *et al*., 2000)
Liaquatpur	6,727	698,985	1,389	290	452	28	619	1,255	134	2001-03 (Kichi *et al*., 2000)
Total	15,594	5,140,284	**59,296**	**13,319**	**21,098**	**3,722**	**21,157**	**56,199**	**3,097**	
			%age	**22.46**	**35.58**	**6.28**	**35.68**	**94.78**	**5.22**	

*According to the 1998 census report (Pakistan Census Bureau)

Table II.- Allelic frequencies and heterozygosities at *ABO* and *Rh* loci.

Population	ABO locus			HWE Test Statistics; p value	Rh locus		Heterozygosity		
	p[A]	q[B]	r[O]		Rh+[D]	Rh-[d]	ABO	Rh	Average
DG Khan	0.1677 ± 0.0029	0.1979 ± 0.0032	0.6345 ± 0.0039	0.920; < 0.05	0.7653 ± 0.0051	0.2349 ± 0.0051	0.5302 ± 0.0032	0.3592 ± 0.0034	0.4451 ± 0.0855
Muzaffargarh	0.1892 ± 0.0027	0.2689 ± 0.0031	0.5419 ± 0.0036	121.88; < 0.0001*	0.7410 ± 0.0045	0.2590 ± 0.0045	0.5938 ± 0.0020	0.3838 ± 0.0028	0.4910 ± 0.1072
Multan	0.1837 ± 0.0056	0.2712 ± 0.0066	0.5451 ± 0.0075	0.365; < 0.05	0.7812 ± 0.0094	0.2188 ± 0.0094	0.5957 ± 0.0043	0.3419 ± 0.0064	0.4691 ± 0.1268
Bahawalpur	0.1423 ± 0.0014	0.2389 ± 0.0017	0.6187 ± 0.0020	107.47; < 0.0001*	0.7879 ± 0.0026	0.2121 ± 0.0026	0.5398 ± 0.0016	0.3343 ± 0.0018	0.4368 ± 0.1027
Liaquatpur	0.1231 ± 0.0064	0.1929 ± 0.0079	0.6840 ± 0.0093	30.86; < 0.0001*	0.6894 ± 0.0128	0.3106 ± 0.0128	0.4799 ± 0.0092	0.4284 ± 0.0067	0.4544 ± 0.0256
Total	0.1566 ± 0.0011	0.2386 ± 0.0013	0.6048 ± 0.0015	176.25; < 0.0001*	0.7715 ± 0.0020	0.2285 ± 0.0020	0.5527 ± 0.0011	0.3526 ± 0.0013	0.4537 ± 0.1002

*Significant deviation from HWE expectations.

To understand the diversity at *ABO* and *Rh* allelic systems, individual and combined heterozygosities were calculated. Heterozygosity was observed to be higher at *ABO* locus ranging from 0.4799±0.0092 in Liaquatpur to 0.5938±0.0020 in Muzaffargarh, while at *Rh* locus it ranged from 0.3343±0.0018 in Bahawalpur to 0.4284±0.0067 in Liaquatpur (Table II). Collectively, the heterozygosity estimates at *ABO* and *Rh* loci were 0.5527±0.0011 and 0.3526±0.0013, respectively, and averaged 0.4537±0.1002 (Table II). These analyses demonstrated that average heterozygosity and gene diversity was maximum in Muzaffargarh (0.4910±0.1072), and minimum in Bahawalpur (0.4368±0.1027) (Table II). It was further established that among the two genetic systems employed in this study, the *ABO* locus had more resolving power.

Heterogeneity between the recruited populations was evident when Z-tests were performed by utilizing the phenotypic proportions at the ABO and Rh systems. No two populations were concordant for all six phenotypes (*i.e.*, A, B, AB, O, Rh+, Rh-). Additionally, pair-wise homogeneity tests with allelic frequencies also revealed marked differences between these populations, and limited systems of similarities were evident. For example, at *ABO* locus, only one homogeneous pair emerged *i.e.*, Multan and Muzaffargarh (χ^2 = 0.79; df. 2; $p<0.05$). At the *Rh* locus, there were two homogeneous pairs *i.e.*, Multan and DG Khan (χ^2 = 2.11; df. 2; $p<0.05$), and Multan and Bahawalpur (χ^2 = 0.48; df.2; $p<0.05$). Heterogeneity was increasing when three or more populations were considered simultaneously. For instance, there was significant evidence against homogeneity (χ^2 = 360.60; df. 4; $p<0.05$) when Multan, Muzaffargarh and DG Khan populations were compared. Heterogeneity was ever more pronounced when all five populations were considered simultaneously.

Degree of differentiation between the Southern Punjab populations was established by gene diversity analyses (Nei, 1975). High degree of differentiation and fragmentation was witnessed in the Southern Punjab populations (*i.e.*, coefficient of gene differentiation, GST = 0.0584; Table III). The total heterozygosity (HT) was higher than the heterozygosity within subpopulations (HS), suggesting that the genetic differentiation was contributed by inter-population variances (Table III). Standard genetic distances and DA distance matrices established high affinities between the populations of Muzaffargarh and Multan districts (Fig. 2). Additionally, the populations of DG Khan and Bahawalpur were also showing close affinities, while Liaquatpur appeared as outlier among the five populations studied (Fig. 2).

Table III.- Gene diversity indices estimated for the Southern Punjab population.

Locus	Genetic indices			
	HT	**HS**	**DST**	**GST**
ABO	0.5998	0.5606	0.0392	0.0653
Rh	0.4112	0.3913	0.0199	0.0483
All	**0.5055**	**0.4759**	**0.0296**	**0.0584**

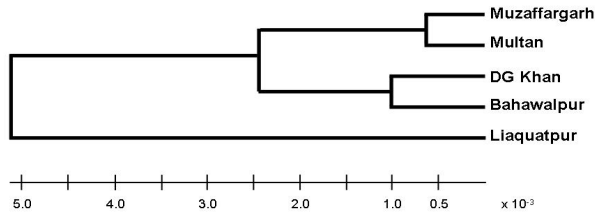

Fig. 2. Dendrogram based upon DA-UPGMA showing the genetic affinities among the five Southern Punjab populations. Two clusters of similarities are evident: Muzaffargarh and Multan; and DG Khan and Bahawalpur, while Liaquatpur appears as an outlier.

DISCUSSION

Southern Punjab in Pakistan is ethnically and linguistically different assemblage of population from upper Punjab/ Punjab. In order to understand the genetic structure of certain Southern Punjab populations we have conducted a preliminary study by employing two polymorphic genetic systems, *i.e.*, *ABO* and *Rh* loci. For these analyses we have included the populations with substantially large samples available (> 1,000 subjects). All Southern Punjab districts with small sample sizes (*i.e.*, Rahim Yar Khan, Lodhran) and the upper Punjab populations were excluded (Shami and Rasmuson, 1994; Mian and Farooq, 1999).

The phenotypic proportions at *ABO* and *Rh* loci in the Southern Punjab populations were significantly different from that of upper Punjab populations ($p<0.0001$) (Rehman *et al*., 2005; Shami and Rasmuson, 1994). The deviations were significant even when averaged phenotypic proportions of all Punjab were considered (excluding cosmopolitan populations of Lahore/ Rawalpindi). This suggested a different composition of allelic frequencies and heterozygosities at studied loci in Southern Punjab populations compared to upper Punjab/Punjab.

Three populations in our data set were depicting significant deviations from the HWE (Table II). Deviations from HWE may result due to the disproportionate phenotypic categories within blood groups. Close inspection of the phenotypic percentage in our data and the output of Z-tests gave strong evidence of this situation, *i.e.,* over- and under-represented phenotypic classes. For instance, in Muzaffargarh population there was over-representation of B blood type while type O was under-represented (Table I). In Bahawalpur and Liaquatpur Districts, AB type was extremely rare (5.63% and 2.02%, respectively). In Liaquatpur, O blood group was the most prevalent (44.56%) and over-represented.

Furthermore, assortative mating systems within a large population lead to isolated, stratified, and endogamous sub-populations which show deviations from HWE (Nei, 1975). In the case of Southern Punjab, at least some part of deviation from HWE could be attributed to the high rate of inbreeding/consanguineous marriages. This view is supported from the current data (*i.e.,* low estimates of HS; Table III), and the recent studies carried out in Southern Punjab. For instance, a genetic epidemiological study in three Districts of Southern Punjab has shown that more than 43% of all marriages were consanguineous/cousin unions (Afzal, 2010). When marriages within relatives/*bradri* were also considered then the proportion of such marriages reached up to 82% (Afzal, 2010). Another independent investigation revealed that close marriages comprised 54% of total marital unions in DG Khan, and ~70% of all marriages were endogamous (Amin-ud-Din, 2009). Additionally, previous studies have also established that Southern Punjab populations maintained highest inbreeding coefficient (*F*) in Pakistan (Shami *et al*., 1994). All these studies support the notion that in the recruited Southern Punjab populations, endogamous marriages might have contributed, at least in part, to the obvious deviations from HWE. Finally, the other factors which could result in deviations from HWE are bias of ascertainment, non-random sampling, and unequal male/female ratios. In conclusion, the HWE deviation may also depict substantial heterogeneity within populations and the existence of sub-populations within the districts.

The evaluation of locus heterozygosities in the studied populations revealed that heterozygosity at sub-population level, HS was much lower than the overall heterozygosity (HT) at the individual and combined *ABO* and *Rh* loci (Table III). Particularly at the *ABO* locus, there was a pronounced loss of heterozygotes (and elevated number of homozygotes), which substantially imparts to the overall depletion of heterozygosity (Table III). The comparison of these results with the previous studies demonstrated that heterozygosity in total and in sub-populations in Southern Punjab was higher than similar estimates for all Punjab (Shami and Rusmoson, 1999). The heterozygosity in Southern Punjab was also significantly higher than the populations of Khyber Pukhtoonkhwa (Ali, 2010), and Bajaour and Mohmand Agecies and Northern Pakistan (Atta-ur-Rahman, 2011). Likewise, absolute gene diversity (DST) and coefficient of differentiation (GST) were more pronounced in Southern Punjab than all Punjab (Table III). Interestingly, in the genetic differentiation of Southern Punjab *ABO* locus contributes more than the *Rh* locus. Contrastingly, in the differentiation of Northern Pakistan populations, the contribution of *Rh* locus was higher than *ABO* locus (Ali, 2010; Atta-ur-Rahman, 2011). These observations may represent that different evolutionary forces are in operation in these populations occupying different geographic locations.

Phylogenetic analysis carried out by employing DA distance matrix showed close genetic affinities between the populations of Muzaffargarh and Multan (Fig. 2). This could be explained by the fact that both Districts are adjoined by an extended

border and they may comprise similar population structures. The similarities could also be explained by the fact that a large number of individuals from Muzaffergarh go Multan for occupational purposes and may also visit various health institutes there. Curiously however, close affinities were also observed between DG Khan and Bahawalpur populations (Fig. 2). This observation is rather surprising as both these Districts are situated widely apart (*i.e.*, 125 km), and do not share any boundary. The linguistic and ethnic assemblages of DG Khan and Bahawalpur populations are distinct and the present observation needs further elucidation through highly polymorphic genetic markers. Liaquatpur population on the other hand, emerged as an outlier among the five populations studied. This also requires further study in order to exclude that it is not due to sampling.

The immunoserological markers of ABO and Rh blood types cannot present a complete picture of the genetic variation within and between populations. For a comprehensive view, a large number of highly polymorphic markers, *i.e.*, microsatellites are required which could identify the potential evolutionary forces shaping the population. However, since ABO and Rh systems have been thoroughly and globally investigated, they remain very relevant for preliminary population genetic studies. The present study is a first attempt to understand the genetic heterogeneity and gene diversity in the Southern Punjab populations by utilizing *ABO* and *Rh* loci. This pilot study would help understand the genetic structure of Southern Punjab population and to find its affinities with other Pakistani populations.

ACKNOWLEDGEMENTS

The participation of volunteer subjects in this study is highly acknowledged. We appreciate the cooperation of MS and hospital staff at the DHQ Hospitals at DG Khan and Muzaffargarh.

REFERENCES

ABDULLAH, A., 1973. *The historical background of Pakistan and its people*. Tanzeem Publishers, Karachi.

AFZAL, M., 2010. *Prevalence of hereditary and congenital anomalies in three districts of Southern Punjab, Pakistan*. M.Phil. dissertation, Department of Animal Sciences, Quaid-i-Azam University, Islamabad, Pakistan.

ALI, S., 2010. *Genetic heterogeneity, gene diversity and clinal analysis at ABO and Rh loci in NWFP population*. M.Phil. dissertation. Department of Animal Sciences, Quaid-i-Azam University, Islamabad, Pakistan.

AMIN-UD-DIN, M., 2009. *A study of certain aspects of human genetics including consanguinity and genetic disorders in the human population of DG Khan*. Ph.D. dissertation. Bahauddin Zakariya University, Multan, Pakistan.

ATTA-UR-REHMAN., 2011. *Gene diversity and clinal analysis at ABO and Rh loci in Bajaur and Mohmand Agencies*. M.Phil. dissertation. Department of Animal Sciences, Quaid-i-Azam University, Islamabad, Pakistan.

GARSTMAN, B.B., 2008. *Biostatistics*. Jones and Bartlett Publishers, Sudbury, Massachusetts, pp. 380-382.

HASSAN, A. AND RAZA, M., 2009. *Migration and small towns in Pakistan*. International Institute for Environment and Development (IIED), London, UK.

KHICHI, Q.K., ALI, S.M.A. AND MALIK, M.S., 2000. Pervalence of ABO and Rh (D) blood groups in Bahawalpur division. *Pak. Ped. J.*, **24:** 1-2.

MATHER, W.B., 1964. *Principles of quantitative genetics*. Burgess Publishing Co, Minnesota. USA.

MIAN, A. AND FAROOQ, A., 1999. Distribution of ABO and Rh blood group alleles in different populations of Southern Punjab, Pakistan. *Anthrop. Anz.*, **57:** 33-39.

MIAN, A. AND BHUTTA, M.A., 1993. Genetic studies on human population of Multan (Pakistan): Distribution of ABO and Rh blood alleles. *Pakistan J. Zool.*, **25:** 149-151.

NEEL, J.V. AND SCHULL, W.J., 1954. *Human heredity*. University of Chicago Press, Chicago.

NEI, M., 1975. *Molecular population genetics and evolution*. North-Holland/American Elsevier, Amsterdam,. pp. 149-158.

REHMAN, A,. KHAN, M.A., ASHRAF, M., MALIK, S.A., SAEED, M.A., RAFIQUE, A. AND ALI, A., 2005. ABO and Rh blood groups. *Profess. med. J.*, **12:** 368-371.

SHAMI, S.A. AND RASMUSON, M., 1994. Genetic heterogeneity and gene diversity in the population of Punjab, Pakistan, based on ABO and Rh(D) blood group frequencies. *Hum. Hered.*, **44:** 214-219.

SHAMI, S.A., GRANT, J.C. AND BITTLES, A.H., 1994. Consanguineous marriages within social/occupational class boundaries in Pakistan. *J. Biosoc. Sci.*, **26:** 91-96.

SILVA, P.J.N., 2002. *ABO estimator*. Faculdade de Ciencias da Universidade de Lisboa, Lisboa, Portugal.

SNEATH, P.H.A. and SOKAL, R.R., 1973. *Numerical taxonomy*. WH Freeman and Company, San Francisco.

(Received 30 July 2011, revised 13 July 2013)